U.S. Predicted Cancer Incidence, 1999: Complete Maps by County and State From Spatial Projection Models

This publication was prepared by:

Linda Williams Pickle, Ph.D.
Eric J. Feuer, Ph.D.
Brenda K. Edwards, Ph.D.

Surveillance Research Program
Division of Cancer Control and Population Sciences
National Cancer Institute
6116 Executive Blvd., Suite 504
Bethesda, Maryland 20892-8317
Fax: 301-480-2046

Suggested citation for the monograph:

Pickle LW, Feuer EJ, Edwards BK. *U.S. Predicted Cancer Incidence, 1999: Complete Maps by County and State From Spatial Projection Models*. NCI Cancer Surveillance Monograph Series, Number 5. Bethesda, MD: National Cancer Institute, 2003. NIH Publication No. 03-5435.

This publication is available on the SEER web site: **http://srab.cancer.gov/incidence**

Acknowledgments

The author wishes to thank the staff at IMS, Inc., for computer support, particularly Timothy McNeel and Mike Barrett, and Marsha Reichman and B. Sue Bell of NCI for their thoughtful comments that clarified the presentation. The advice of Bradley Carlin, University of Minnesota, greatly shortened the time needed to implement the hierarchical models. Thanks also to Jim Pearson for insightful discussions and novel ideas for the graphs and tables. The designs of the micromap plots and maps are the result of ongoing collaborations with Daniel Carr, George Mason University, and Cynthia Brewer and Alan MacEachren, The Pennsylvania State University Department of Geography.

Table of Contents

State Maps

County Maps

Micromap Plots

Foreword

While NCI has published U.S. cancer mortality maps since 1974, this monograph presents, for the first time, state- and county-level maps of estimated cancer incidence. These estimates are based on a statistical modeling of county-level demographic and lifestyle characteristics, in addition to data from the Surveillance, Epidemiology, and End Results (SEER) Program. This new ability to map both incidence and mortality enables us to explore issues including survival and effects of screening. Data are included for total cancers; for the four most frequent cancers—lung and bronchus, colon and rectum, prostate, and breast; and for all other cancer sites combined.

The maps presented here represent a qualitative advance in their use of state- and county-level sociodemographic and lifestyle data for estimation. Previous estimates of cancer incidence by state have assumed that the ratio of each state's incidence to mortality is the same as that found for the combined SEER registries, an assumption we know is not justified in all cases.

These data fill gaps where state cancer registries have not yet reached the level of complete reporting required for inclusion in the United States Cancer Statistics (USCS). However, rapid progress is being made toward that goal, but even after it is reached these data will prove valuable on both national and state levels.

From a national perspective, the maps included in this monograph allow examination of the geographic distribution of cancer incidence across the country and of the magnitude of differences among states. They show higher predicted incidence rates for lung cancer in states in the Southeast, for colorectal cancer in midwestern states, and for all cancers combined in northeastern states. A greater range of predicted incidence rates among states is observed for lung and colorectal cancers than for other cancers.

Smoothed maps of county-level incidence allow us to see the differences among geographic regions other than by state only. This is important because using administrative boundaries, such as state borders, may not be the most accurate or meaningful method of tabulating differences in cancer rates. For example, the high predicted female lung cancer incidence rates for counties along the northern Pacific coast are clearly visible in smoothed county maps. Smoothed county-level maps of cancer incidence may also allow correlation with geologic data or environmental data of other types. County-level maps may allow those with community-level knowledge to see correlations between local conditions and cancer incidence patterns.

From the perspective of individual states, these data offer the ability to utilize county-level data to provide estimates of the numbers of new cancer cases expected at the beginning of the data collection year. Importantly, this will allow cancer control specialists to target interventions to specific areas by using these data in conjunction with information from various state programs (e.g., screening and early detection) and with demographic characteristics including income, race/ethnicity, medical insurance, etc. These data are also useful for quality control both for states that are in the process of improving their cancer registries and for states where the variation in cancer incidence from the national levels is sufficiently great that predictions are needed that emphasize local conditions rather than the national average.

We hope that the presentation of this data will excite and stimulate researchers and those involved in cancer surveillance, control, and prevention activities to utilize this novel approach to further reduce the cancer burden in America.

Robert T. Croyle, Ph.D.
Director
Division of Cancer Control and Population Sciences
National Cancer Institute
National Institutes of Health

Introduction

The primary source of information about cancer incidence in the United States for the past 30 years has been the Surveillance, Epidemiology, and End Results (SEER) program of the National Cancer Institute (NCI) (see **http://seer.cancer.gov**). With the most recent expansion of the SEER Program, these population-based cancer registries now cover approximately 26% of the U.S. population. However, even in combination with high-quality registries from additional states funded by the CDC National Program for Cancer Registries (NPCR), gaps in the collection of data prevent the calculation of cancer incidence statistics for many states, for regions, and for the U.S. as a whole (USCS 2002). In this report, we present the results of a statistical model that predicts the number of new cases and incidence rates for the major types of cancer for every U.S. state and county.

Estimates of the numbers of new cancer cases and rates expected in an area are useful for cancer surveillance, cancer control, health resource planning, and quality control activities. Geographic targeting of cancer-related activities to local areas with the most need has been shown to greatly improve their effectiveness (Kerner et al. 1988). For example, scarce health department resources for cancer prevention programs can be allocated to locations of the greatest need. In addition, with more accurate estimates of expected cases, the cancer registrar can monitor the cumulative number of cases found throughout the year to judge the degree of completeness of data collection and to identify locations with unexpectedly high counts that may require further investigation. This independent source of expected case counts can also provide an objective means of deriving a completeness index for certification purposes (Tucker and Howe 2001). Finally, by providing a complete set of predicted rates for each state and region, state registries may compare their cancer experience with that of neighboring states.

Currently, the only source of complete estimates of expected case counts and rates by state is the American Cancer Society's annual report *Cancer Facts and Figures* (ACS 2003). These figures are the result of both spatial and temporal projections. For each year, they compute the estimated number of new cancer cases and rates for each state using the cancer incidence rates and cancer incidence-mortality ratios aggregated across all SEER registries and each state's mortality and population for that year (Wingo et al. 1998). Then these state estimates are projected ahead in time several years using a time series model to provide a set of expected numbers for the next calendar year. This method has the potential for improvement, particularly for cancers whose rates vary by geographic area.

The results presented here are computed by a spatial projection model that predicts the number of cases in each county based on the sociodemographic and lifestyle profile for that county. The utility of this approach was demonstrated by an analysis of SEER breast cancer incidence using similar methods with a simpler model (Frey et al. 1993). By the inclusion of these additional cancer risk factors and by allowing the cancer rates to vary by geographic area, these results should form a better basis for the temporal projection of state data. We are currently working to extend this method to project these spatial estimates ahead in time to provide state estimates for the upcoming calendar year. Collaborations are underway with the ACS to incorporate these improvements into their annual report.

The purpose of this report is to present complete county and state maps and tables of rates and case counts for 1999 estimated by these new statistical models. Numbers of cases and rates are shown with and without adjustment for reporting delay (Clegg et al. 2002) and figures reported in the recently published *United States Cancer Statistics: 1999 Incidence* are shown for comparison (USCS 2002). Differences between the predicted cancer incidence figures and those reported in USCS cannot be ascribed to any particular source without further exploration, and we urge readers to take a systematic approach in exploring them. This monograph demonstrates that this new method can successfully be used to predict cancer incidence. Not only can this method fill in the current gaps in cancer data collection, but even when all U.S. states collect their own cancer data, it can provide a baseline expectation for the cancer incidence in each state for the coming year.